SEMINAR

Introduction to Instrumental Variable Analysis

Shotaro Aso¹, Hideo Yasunaga²

¹ Department of Biostatistics and Bioinformatics, Graduate School of Medicine, The University of Tokyo

² Department of Clinical Epidemiology and Health Economics, School of Public Health, The University of Tokyo

ABSTRACT

In theory, instrumental variable (IV) analysis, like randomized controlled trials, can adjust for measured and unmeasured confounders. IVs need to meet the following three conditions: (i) they are associated with treatment assignment; (ii) they have no direct association with the outcome and are associated with the outcome exclusively through the treatment; and (iii) they are not associated with any of the measured confounders. Studies have presented several types of IV, including preferences of the facility or physician, differential distance, and days of the week. Two types of estimation method have been introduced: two-stage least squares and two-stage residual inclusion. The assumption of monotonicity limits the generalizability of estimates of causal effects in IV analysis because the target population of IV analysis is “compliers” (those who always comply with the assigned treatment). IV analysis using two or more IVs is feasible but requires the overidentifying restriction test. Despite several limitations, IV analysis is a feasible option that may be used for causal inference in comparative effectiveness studies using retrospective observational data.

KEY WORDS

instrumental variable analysis, two-stage least squares, two-stage residual inclusion, monotonicity, causal inference

INTRODUCTION

Propensity score analysis cannot adjust for unmeasured confounders because it is conducted with the assumption that there are no unmeasured confounders [1, 2]. Conversely, instrumental variable (IV) analysis, in theory, is able to adjust for both measured and unmeasured confounders. IV analysis was originally introduced in economics as a tool for causal inference [3]. Its use in clinical studies remain uncommon. Although IV analysis is not an all-around method and despite several limitations, it can be seen as a feasible option for addressing confounding bias in clinical studies using retrospective data. This article explains the basis and limitations of IV analysis.

MECHANISM

IV analysis resembles a randomized controlled trial. In randomized controlled trials, eligible patients are randomly assigned to the treatment group or the control group

(**Fig. 1**). Random assignment equalizes patients’ backgrounds across the two groups and prevents bias caused by measured and unmeasured confounders. Random assignment involves the following three conditions: (i) it is directly associated with treatment allocation; (ii) it is associated with the outcome exclusively through the treatment and has no direct association with the outcome; and (iii) it is not associated with patient background factors (measured confounders). IV analysis is similar to randomized controlled trials in that, in IV analysis, the IVs play the role of random assignment in randomized controlled trials [4] (**Fig. 2**). IVs needs to meet the same three assumptions introduced above: (i) IVs should be strongly associated with treatment allocation (**Fig. 3**); (ii) IVs should not be directly associated with the outcomes but should be associated with the outcomes only through the treatment (**Fig. 4**); and (iii) IVs should not be associated with measured confounders (**Fig. 5**). IV analysis introduces pseudo-randomization and can adjust for both measured and unmeasured confounders.

Assumption (i) can be quantitatively assessed using the *F*-statistic [5]. The *F*-statistic is an indicator of how the IV is associated with treatment allocation and is estimated in a multivariable regression where the dependent variable is the treatment variable and the independent variables are the IV and the measured confounders. When the *F*-statistic is >10 , the IV is strongly associated with treatment allocation. When the *F*-statistic is ≤ 10 , the IV is not associated with treatment allocation; such an IV is called a “weak” IV. IV analysis using weak IVs can produce misleading results.

Assumption (ii) cannot be quantitatively assessed [6]. Researchers must clinically judge whether the condition of no direct association between the IV and the outcome is met. In studies using IV analysis, researchers should compare outcomes between individuals in different categories of the IV [7, 8].

Assumption (iii) also cannot be quantitatively assessed [6]. However, reporting the associations between the measured confounders and the IV is recommended. If the IV is a continuous variable, it should be categorized into two groups using an appropriate cut-off value, and researchers should then compare measured confounders between two resulting IV categories. The distributions of the measured confounders are not always well-balanced between these groups [9, 10].

TYPES OF INSTRUMENTAL VARIABLE

Several types of IV have been presented in comparative effectiveness studies.

Facility or Physician Treatment Rates

Facility or physician treatment rates are the best-known

Fig. 1 Randomized controlled trial

Fig. 2 Instrumental variable analysis

Fig. 3 Condition (i) Instruments are associated with assignment of treatment

Fig. 4 Condition (ii) Instruments are not directly associated with outcome

Fig. 5 Condition (iii) Instruments are not associated with measured confounders

type of IV. A previous study showed that studies using facility or physician treatment rates accounted for approximately 40% of all epidemiological studies using IVs [11]. Facility and physician treatment rates successfully meet the three conditions described above. Patients admitted to a hospital that frequently chooses the treatment tend to be allocated to the treatment group [12]. Thus, facility and physician treatment rates meet assumption (i). Because facility and physician treatment rates do not affect the outcome or the confounders, they also meet assumptions (ii) and (iii).

Differential Distance

The definition of differential distance (DD) is the difference between the distance from a patient's home to a hospital that frequently selects the treatment (d_1) and the distance from the patient's home to the nearest hospital (d_2 ; $DD = d_1 - d_2$; Fig. 6) [9]. When the hospital nearest to the patient's home is a hospital that frequently selects the treatment, $DD = 0$. When DD approaches zero, patients tend to receive the treatment. Conversely when DD increases, patients tend not to receive the treatment. This mechanism explains how DD meets assumption (i). Because DD does not affect the outcome or the measured confounders, DD also meets assumptions (ii) and (iii).

Dates

Dates are a frequently used type of IV. For example, the day of the week of patient admission to a hospital can be used as an IV [13]. Hospitals are less likely to perform certain examinations or treatments over the weekend.

Such a situation can result in less opportunity to be allocated to the treatment group on weekends. Thus, as an IV, date meets assumption (i). Generally, dates are not associated with the outcome or patient characteristics, and dates therefore also meet assumptions (ii) and (iii).

ESTIMATION METHODS

IV analysis generally involves a two-stage modeling approach to estimate treatment effects. In the first stage, the association of the IV with the actual treatment allocation is estimated. In the second stage, the outcomes are compared in terms of the predicted probability of receiving the treatment rather than the actual receipt of the treatment. Here, we describe two representative estimation methods.

Two-stage Least Squares

Two-stage least squares is the best-known method for IV analysis and has traditionally been adopted in studies using IVs [14]. Two-stage least squares is often used when the outcomes are continuous variables. Because linear regression is used in two-stage least squares, the associations between the IV and the treatment and between the treatment and the outcomes are assumed to be linear. The two-stage least squares procedure is as follows:

- (i) Linear regression is performed to estimate the effect of the IV on treatment allocation, with adjustment for the measured confounders (first stage).
- (ii) The probability of receiving the treatment for each patient is predicted based on the linear regression coefficient.

cient estimating the effect of the IV on the treatment.

(iii) Linear regression is conducted with independent variables including the measured confounders and the predicted probability of receiving the treatment (rather than the actual receipt of the treatment) and the outcome as the dependent variable (second stage). The coefficient of treatment allocation represents the marginal effect of the treatment on the outcome.

Two-stage Residual Inclusion

Two-stage residual inclusion can be used when the outcome is either a binary variable or a continuous variable. Two-stage residual inclusion can accommodate both linear and non-linear associations between the IV and the treatment and between the treatment and the outcome [15]. A previous study applied Cox proportional hazard models in the second stage [9]. Two-stage residual inclusion uses the “residual”—that is, the difference between the actual receipt of the treatment (binary value of 0 or 1) and the predicted probability of receiving the treatment (continuous value from 0 to 1). The two-stage residual inclusion procedure is as follows:

(i) Linear or non-linear regression is performed to estimate the effect of the IV on treatment allocation, with adjustment for confounders (first stage).

(ii) The residual is calculated for each patient based on the coefficient of the regression for the effect of the IV on treatment allocation.

(iii) Linear or non-linear regression is conducted with independent variables including the measured confounders, the residual calculated above, and the actual receipt of the treatment and the outcome as the dependent variable (second stage). The coefficient of treatment allocation denotes the marginal effect of the treatment on the outcome.

MONOTONICITY

The IV analysis assumption of monotonicity limits the generalizability of estimates of causal effects in IV analysis [16]. With a binary IV and a binary treatment, if we could observe the actual assignment of treatment and the counterfactual assignment of treatment, we would have

four different subgroups (**Table 1**). In reality, we can observe only the actual assignment of treatment, and we cannot distinguish between these four subgroups.

“Always-takers” always receive the treatment, independent of whether they are assigned to it. “Never-takers” never receive the treatment, independent of whether they are assigned to it. IV analysis cannot estimate the causal effect of always-takers or never-takers because assignment of treatment for these individuals does not depend on their value on the IV. “Compliers” always comply with the assigned treatment, meaning that they receive the treatment only when they are assigned to it and that they do not take it when not assigned to it. “Defiers” are those who receive the treatment when they are not assigned to it or do not receive the treatment when they are assigned to it. In IV analysis, it is assumed that there are no defiers because the IV is associated with the assignment of treatment. The target population for IV analysis is thus not the entire population; rather, it is only the population of compliers [6, 17]. The average treatment effect in IV analysis is called the “complier average treatment effect” or the “local average treatment effect” [18].

USING TWO OR MORE INSTRUMENTAL VARIABLES

It is possible to use two or more IVs to estimate the effect of one treatment. The *F*-statistic is higher in IV analyses using two or more IVs than in those using one IV. When we use two or more IVs, they must not be endogenous variables. An endogenous variable is synonymous with a dependent variable that correlates with other factors in the model. When IVs are endogenous variables, they do not meet assumptions (ii) and (iii). Overidentifying restriction tests should be conducted when two or more IVs are used [5].

REPORTING

A previous report showed that many studies using IV analysis did not report enough information to determine whether the inferences the authors drew were supported

Table 1 Monotonicity			
		Instrument = 0	
		Treatment = 1	Treatment = 0
Instrument = 1	Treatment = 1	Always-taker	Complier
	Treatment = 0	Defier	Never-taker

by their evidence [11]. The report proposed a checklist of information to guide IV analysis, arguing that, when IV analysis is conducted, researchers should ensure that the following elements are accomplished in reporting on the study: (1) state which population target parameter the study aims to estimate; (2) report the associations of the IVs and treatment using the F -statistic; (3) report and test the associations of observed potential confounders with both the treatment and the IVs; (4) report the test of overidentifying restriction with multiple IVs; (5) report a tabulation of the frequencies of all combinations of IVs, treatment, and outcome; and (6) always use robust or bootstrapped standard errors and take the clustering of study participants into account where necessary when using generalized linear models with binary outcomes.

LIMITATIONS

In theory, IV analysis adjusts for unmeasured confounders.

However, this method has several limitations. First, few IVs satisfy the three required assumptions. It is therefore difficult to find appropriate IVs, which is a serious limitation. Weak instruments with an F -statistic value <10 lead to misleading results in IV analysis [5]. Second, there is no approach for proving that an IV meets assumption (ii). Instrument–outcome confounders, which affect both the IV and the outcome or mediate the effect of the IV on the outcome, may exist [19].

CONCLUSION

IV analysis can, in theory, adjust for measured and unmeasured confounders, similar to randomized controlled trials. IVs must meet three assumptions. When researchers use IV analysis, they should pay attention to the assumptions and the limitations of this method.

REFERENCES

1. Austin PC, Mamdani MM, Stukel TA. The Use of the Propensity Score for Estimating Treatment Effects: Administrative versus Clinical Data. *Stat Med* 2005;24:1563–78.
2. Austin PC, Grootendorst P, Anderson GM. A Comparison of the Ability of Different Propensity Score Models to Balance Measured Variables between Treated and Untreated Subjects: A Monte Carlo Study. *Stat Med* 2007;26:734–53.
3. Angrist JD. Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *Am Econ Rev* 1990;80:313–36.
4. Greenland S. An Introduction to Instrumental Variables for Epidemiologists. *Int J Epidemiol* 2000;29:722–9.
5. Staiger D, Stock J. Instrumental Variables Regression with Weak Instruments. *Econometrica* 1997;65:557–86.
6. Hernán MA, Robins JM. Instruments for Causal Inference: An Epidemiologist's Dream? *Epidemiology* 2006;17:360–72.
7. Valley TS, Sjöding MW, Ryan AM, Iwashyna TJ, Cooke CR. Association of Intensive Care Unit Admission with Mortality among Older Patients with Pneumonia. *JAMA* 2015;314:1272.
8. Slaughter JL, Reagan PB, Newman TB, Klebanoff MA. Comparative Effectiveness of Nonsteroidal Anti-inflammatory Drug Treatment vs No Treatment for Patent Ductus Arteriosus in Preterm Infants. *JAMA Pediatr* 2017;171:1–7.
9. Tan H-J, Norton EC, Ye Z, Hafez KS, Gore JL, Miller DC. Long-term Survival Following Partial vs Radical Nephrectomy among Older Patients with Early-stage Kidney Cancer. *JAMA* 2012;307:1629–35.
10. Valley TS, Iwashyna TJ, Cooke CR, Sinha SS, Ryan AM, Yeh RW, et al. Intensive Care Use and Mortality among Patients with ST Elevation Myocardial Infarction: Retrospective Cohort Study. *BMJ* 2019;365:l1927.
11. Davies NM, Davey Smith G, Windmeijer F, Martin RM. Issues in the Reporting and Conduct of Instrumental Variable Studies: A Systematic Review. *Epidemiology* 2013;24:363–9.
12. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating Short-term Drug Effects Using a Physician-specific Prescribing Preference as an Instrumental Variable. *Epidemiology* 2006;17:268–75.
13. Sandhu AT, Heidenreich PA, Bhattacharya J, Bundorf MK. Cardiovascular Testing and Clinical Outcomes in Emergency Department Patients with Chest Pain. *JAMA Intern Med* 2017;177:1175–82.
14. Uddin MJ, Groenwold RH, de Boer T, Belitser SV, Roes KC, Klungel OH. Instrumental Variable Analysis in Epidemiologic Studies: An Overview of the Estimation Methods. *Pharm Anal Acta* 2015;6:4.
15. Terza JV, Basu A, Rathouz PJ. Two-stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *J Health Econ* 2008;27:531–43.
16. Swanson SA, Miller M, Robins JM, Hernán MA. Definition and Evaluation of the Monotonicity Condition for Preference-based Instruments. *Epidemiology* 2015;26:414–20.
17. Harris KM, Remler DK. Who is the Marginal Patient? Understanding Instrumental Variables Estimates of Treatment Effects. *Health Serv Res* 1998;33:1337–60.
18. Angrist JD, Krueger AB. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *J Econ Perspect* 2001;15:69–85.
19. Garabedian LF, Chu P, Toh S, Zaslavsky AM, Soumerai SB. Potential Bias of Instrumental Variable Analyses for Observational Comparative Effectiveness Research. *Ann Intern Med* 2014;161:131–8.